

Zur Evaluation von Mensch-Roboter Interaktionen (MRI) - ein methodischer Beitrag aus soziologischer Perspektive

Compagna, Diego; Marquardt, Manuela

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:

Compagna, D., & Marquardt, M. (2015). *Zur Evaluation von Mensch-Roboter Interaktionen (MRI) - ein methodischer Beitrag aus soziologischer Perspektive*. (Working Papers kultur- und techniksoziologische Studien, 03/2015). Duisburg: Universität Duisburg-Essen Campus Duisburg, Fak. für Gesellschaftswissenschaften, Institut für Soziologie. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49288-7>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.



Working Papers
kultur- und techniksoziologische Studien

Volume 8 (3)
no 03/2015

Herausgeber:
Diego Compagna, Stefan Derpmann und Manuela Marquardt
Layout:
Vera Keyzers

Kontakt:
diego.compagna@gmail.com
stefan.derpmann@gmail.com
manuela.marquardt@gmx.de

Ein Verzeichnis aller Beiträge befindet sich hier:
www.uni-due.de/wpkts

ISSN 1866-3877
(Working Papers kultur- und techniksoziologische Studien)

Working Papers kultur- und techniksoziologische Studien - Copyright

This online working paper may be cited or briefly quoted in line with the usual academic conventions. You may also download them for your own personal use. This paper must not be published elsewhere (e.g. to mailing lists, bulletin boards etc.) without the author's explicit permission.

Please note that if you copy this paper you must:

- include this copyright note
- not use the paper for commercial purposes or gain in any way

You should observe the conventions of academic citation in a version of the following form:

Author (Year): Title. In: Working Papers kultur- und techniksoziologische Studien (no xx/Year). Eds.: Diego Compagna / Stefan Derpmann / Manuela Marquardt, University Duisburg-Essen, Germany. www.uni-due.de/wpkts (dd.mm.yyyy)

Working Papers kultur- und techniksoziologische Studien - Copyright

Das vorliegende Working Paper kann entsprechend der üblichen akademischen Regeln zitiert werden. Es kann für den persönlichen Gebrauch auch lokal gespeichert werden. Es darf nicht anderweitig publiziert oder verteilt werden (z.B. in Mailinglisten) ohne die ausdrückliche Erlaubnis des/der Autors/in.

Sollte dieses Paper ausgedruckt oder kopiert werden:

- Müssen diese Copyright Informationen enthalten sein
- Darf es nicht für kommerzielle Zwecke verwendet werden

Es sollten die allgemein üblichen Zitationsregeln befolgt werden, bspw. in dieser oder einer ähnlichen Form:

Autor/in (Jahr): Titel. In: Working Papers kultur- und techniksoziologische Studien (no xx/Jahr). Hrsg.: Diego Compagna / Stefan Derpmann / Manuela Marquardt, Universität Duisburg-Essen, Deutschland. www.uni-due.de/wpkts (tt.mm.jjjj)

Vorwort

Eine soziologische Betrachtung von Technik zeichnet sich unter anderem dadurch aus, dass das Bedingungsverhältnis zwischen den technischen Artefakten und den sozialen Kontexten, in die jene eingebettet sind, als ein interdependentes – zu beiden Seiten hin gleichermaßen konstitutives – angesehen wird. Diesem Wesenszug soziologischer Perspektiven auf Technik trägt der Titel dieser Reihe Rechnung, insofern von einer soziokulturellen Einfärbung von Technik sowie – vice versa – eines Abfärbens von technikhärenten Merkmalen auf das Soziale auszugehen ist. Darüber hinaus schieben sich zwischen den vielfältigen Kontexten der Forschung, Entwicklung, Herstellung, Gewährleistung und Nutzung zusätzliche Unschärfen ein, die den unterschiedlichen Schwerpunktsetzungen und Orientierungen dieser Kontexte geschuldet sind: In einer hochgradig ausdifferenzierten Gesellschaft ist das Verhältnis von Sozialem und Technik durch je spezifischen Ent- und Rückbettungsdynamiken gekennzeichnet.

Die Reihe Working Papers kultur- und techniksoziologische Studien (WPktS) bietet eine Plattform für den niederschweligen Austausch mit Kolleg_innen und steht Wissenschaftler_innen und Student_innen aller Universitäten, Fachrichtungen und Institute für die Veröffentlichung ihrer Forschungs- und Qualifikationsarbeiten offen. Der thematische Rahmen ist hierfür mit Absicht breit gewählt und kann mit verschiedensten Darstellungsformen – vom Essay über die Forschungsskizze bis zum Aufsatz – bearbeitet werden.

Die Reihe WPktS erscheint seit 2008; jede Ausgabe kann Online (www.uni-due.de/wpkts) als PDF-Dokument abgerufen werden.

Die Herausgeber

Berlin und Essen, im April 2015

Zur Evaluation von Mensch-Roboter Interaktionen (MRI) – ein methodischer Beitrag aus soziologischer Perspektive¹

Diego Compagna & Manuela Marquardt

Dr. phil. Diego Compagna / diego.compagna@tu-berlin.de

Manuela Marquardt (cand. M.A.) / manuela.marquardt@gmx.de

Keywords

Krisenexperimente, Rahmenanalyse, Ethnomethodologie, Symbolischer Interaktionismus, Soziale Robotik

Abstract

Trotz ihrer zentralen Bedeutung für die Weiterentwicklung von Mensch-Roboter Interaktionen (MRI), stellt die Evaluation der Interaktion zwischen Menschen und sozialen Robotern nach wie vor eine der wichtigsten methodologischen Herausforderungen dar. Dieser Beitrag stellt auf der Basis von soziologischen Interaktionsmodellen eine theoriegeleitete Methode vor, die dazu in der Lage sein soll, wichtige Aspekte einer gelungenen Interaktionserfahrung zu identifizieren und damit die Qualität der Interaktion zu evaluieren. Sie eignet sich besonders gut für experimentelle Forschung in Laboren und Entwicklungsumgebungen. Im Kern basiert die Methode auf der Anwendung von Harold Gafinkel's Krisenexperimenten, kombiniert mit Erving Goffman's Rahmenanalyse. Den theoretischen Hintergrund der Evaluation bildet eine soziologische Definition sozialer Interaktion aus Theorien des Symbolischen Interaktionismus Paradigmas.

1. Einleitung

Im Bereich der MRI stehen zwei Fragen im Zentrum des Interesses: Erstens, welche Einflussfaktoren bestimmen, ob eine Interaktionserfahrung als gelungen bewertet wird? Zweitens, in welchem Ausmaß und in welchen Situationen werden Interaktionsmodelle präferiert, die sich an konventionellen (sozialen) Interaktionserfahrungen mit anderen Menschen orientieren (etwa im Vergleich zu typischen Mensch-Maschine Interaktionen). Beide Fragen sind eng miteinander verflochten – bezieht sich die erste Frage auf die Bewertung der Qualität der Interaktion, zielt die zweite Frage auf die womöglich² situative Variabilität der Bewertungsmaßstäbe ab. Ziel dieses Beitrags ist es, theoretisch und empirisch fundierte konzeptuelle Überlegungen für eine Methode zu präsentieren, die in der Lage ist, beide Fragen gleichermaßen zu adressieren. Sie basiert im Kern auf dem Prinzip der von Harold Garfinkel

1 Die diesem Beitrag zugrundeliegende Forschung hat im Rahmen einer interdisziplinären Nachwuchsforschungsgruppe stattgefunden, die vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wird.

2 Womöglich – da es auch sein kann, situationsübergreifende Qualitätsmerkmale für gelungene Interaktionserfahrungen zu identifizieren.

entlehnten Krisenexperimente zur Identifikation der Grenzen des (sozialen) Interaktionspielraumes, angereichert mit einer von Erving Goffman inspirierten Rahmenanalyse, um der situativen Variabilität Rechnung zu tragen. Der Beitrag ist wie bereits erwähnt konzeptueller Natur, da die präsentierte Methode in dieser Form noch nicht angewandt wurde – es existieren allerdings empirische Befunde, die für die Eignung der Methode sprechen und gleichzeitig Hinweise für spezifische Konstellationen von Rahmungen und den Umgang mit induzierten Krisen geben – der, wie später zu zeigen sein wird, als Maßstab für die Bewertung der Qualität der Interaktion herangezogen werden kann.

Hierfür wird zunächst der Stand der Forschung (2.) im Bereich der Evaluation von MRI präsentiert, um anschließend die theoretischen Grundlagen (3.) der vorgeschlagenen Methode vorzustellen. Dabei wird näher auf den Begriff symbolisch vermittelter Interaktion, die Ethnomethodologie und die Rahmenanalyse eingegangen, um auf der Basis dieser genuin soziologischen Konzepte und Theorien eine generische Methode vorzuschlagen (4.). Dieser konzeptuelle Teil wird durch den Vergleich zweier empirischer Fallstudien (5.) unterfüttert. Abschließend erfolgt eine Zusammenfassung und kritische Reflexion (6.) der Eignung der hier präsentierten Methode zur Qualitätsevaluation von MRI.

2. Stand der Forschung der (soziologischen) Evaluation von Mensch-Roboter Interaktionen

Die MRI Forschung ist nach wie vor verhaftet in einer psychologischen, und damit – soziologisch betrachtet – methodologisch individualistische³ Sichtweise (vgl. etwa Wykowska et al. 2014, Feil-Seifer et al. 2007, Kahn et al. 2007 und die meisten der Paper in Herrmann 2013).

Soziologisch inspirierte Ansätze, basierend auf soziologischen Modellen, Definitionen und sozialen Handlungs- und Interaktionstheorien, gibt es äußerst wenige (vgl. etwa Burghart/Haeussling 2005, Compagna/Muhl 2012). Burghart und Haeussling entwerfen einen netzwerkbasierten Ansatz, der den operationalen Prozessen von Robotern bzw. Maschinen

³ Diese zeichnet sich dadurch aus, dass soziale Phänomene auf das (absichtsvolle) Handeln einzelner Individuen zurückgeführt werden.

eine direkte soziale Effektivität zuschreibt und evaluieren damit Mensch-Roboter Kooperationen. Die Kooperation wird anhand von vier Ebenen und deren Unterebenen klassifiziert, die soziologische und technische Parameter berücksichtigen. Hierzu gehören der Interaktionskontext (Level A), die Kooperation, verstanden als interagierendes Netzwerk (Level B), die Aktivität und Intentionen der beteiligten Akteure (Level C) und zuletzt die non-verbalen Handlungen und Emotionen (Level D). Eine intuitive Interaktion steht für sie für die Qualität der Kooperation.

Interaktionell-inspirierte Forschungen wurden im Bereich der MRI von der interaktionellen Linguistik proklamiert und vorangetrieben, die auf ethnomethodologischer Konversationsanalyse beruht, welche in engem Zusammenhang mit Garfinkel's Ethnomethodologie steht (vgl. Pitsch et al. 2012, 2013, Pitsch/Wrede 2014). Befunde zu lernfähigen Robotern betonen hier die interaktionelle Dimension des Lernens, die bspw. die Fähigkeit des Sprecherwechsels (turn-taking), der Herstellung gemeinsamer Aufmerksamkeit (joint attention) und die Notwendigkeit geeigneter Feedback Mechanismen (grounding) einschließt (vgl. Wrede et al. 2009).

All diese Ansätze sollten für zukünftige Forschung berücksichtigt werden, um einen übergreifenden, soziologisch inspirierten und interdisziplinär anschlussfähigen Analyserahmen für empirische Studien innerhalb der MRI zu gestalten.

Während die bisherige Forschung zum Großteil auf der Annahme basiert, dass Interaktion in irgendeiner Weise Resultat zweier monolithischer Köpfe ist, die autonom in der Lage sind, konsistente Bedeutungen der Welt herzustellen und ihre Annahmen von Zeit zu Zeit mit anderen abzugleichen, soll die hier präsentierte Evaluationsmethode auf dem Verständnis symbolisch vermittelter Interaktion aufbauen. Dazu werden im Folgenden die theoretischen Grundlagen der hier präsentierten Methode vorgestellt.

3. Theoretische Grundlagen

Der von George Herbert Mead inspirierte symbolische Interaktionismus, Harold Garfinkel's Ethnomethodologie und Erving Goffman's Rahmenanalyse teilen – neben der Kritik an Talcott Parson's Strukturfunktionalismus – die Grundannahme, dass soziale Ordnung von den

Individuen selbst geschaffen wird. Der Fokus liegt hierbei auf symbolischen Interaktionsprozessen, denen sich interpretativ genähert wird. Im Folgenden soll nun knapp auf diejenigen theoretischen Grundlagen der drei Autoren eingegangen werden, die für die darzustellende Methode von Bedeutung sind.

George Herbert Mead entwirft in „Geist, Identität und Gesellschaft“ einen Handlungsbegriff, bei dem die Bedeutung eines Symbols intersubjektiv in der sozialen Interaktion ausgehandelt wird und damit stets von der Reaktion des Anderen (alter) auf eine Aktion von ego abhängt (vgl. Mead 1934/2013). Die Bedeutung konstituiert sich ex post. Soziale Realität ist das Resultat einer erfolgreichen Interaktion zwischen mindestens zwei Entitäten. Dieses von Mead vorgedachte Grundmodell findet sich auch im Sozialkonstruktivismus, der dann verstärkt ein Augenmerk auf die Habitualisierung und Institutionalisierung der zuvor ausgehandelten Symbolbedeutungen legt, die somit erwartbar werden (vgl. Berger/Luckmann 1966/2003). Erwartungen sind somit die Folge von ausgehandelten Bedeutungen. Erving Goffman's Rahmenanalyse (vgl. Goffman 1974) enthält nun Annahmen darüber, wie jede soziale Interaktion von situierten (kontextuellen) Erwartungen durchzogen ist, während Harold Garfinkel's Ethnomethodologie (vgl. Garfinkel 1967) in erster Linie diejenigen Mechanismen in den Fokus nimmt, die genutzt werden – bzw. sich unter den interagierenden Entitäten gemeinhin als brauchbar erwiesen haben – um eine Angleichung der erwarteten Erwartungen auf beiden Seiten auszuhandeln. Garfinkel nimmt dabei eine extreme Position hinsichtlich der Wichtigkeit des Subjekts und seiner oder ihrer Performanz in der Etablierung eines stabilen sozialen Referenzsystems ein. Soziale Akteure müssen in jeder Interaktion ihre Annahmen und ihr Verständnis bezüglich der entsprechenden Situation nachjustieren. Dies tun sie mithilfe von Methoden, mit denen sie sich gegenseitig anzeigen, wie das, was sie tun, zu verstehen ist.

Goffman's Ansatz liegt hinsichtlich der Relevanz der Rahmung hingegen die Annahme zugrunde, dass ein gewisses Maß an Bedeutung bereits vor der Interaktion als fest etabliert gelten kann. Auch wenn die Bedeutung von Handlungen, Worten und anderen Symbolen auf der gegenseitigen Bestätigung durch die Akteure beruht, liegt die eigentliche Mühe darin, die kulturspezifischen Codes zu lernen und korrekt zu interpretieren.

Ogleich sich die beiden Theorien hinsichtlich der Basisannahmen unterscheiden, weisen sie bei allgemeinerer Betrachtung zwei wichtige Ähnlichkeiten auf. Beide betrachten soziale Realität als Ergebnis der Interpretation und Bestätigung (oder Etablierung) der Bedeutungen von Handlungen und Symbolen. Und beide beziehen eine konstruktivistische Sichtweise auf Sozialität. Damit sind sie – obgleich für die Autoren lediglich Mensch-Mensch Interaktionen in Frage kamen – in dieser Hinsicht gut auf MRI Settings anwendbar.

Denn letztlich ist es ein kontingentes Phänomen, dass heutzutage in vielen Kulturen Menschen die einzigen sind, die als soziale Akteure in Frage kommen (vgl. Lindemann 1999, Baecker 2011). Im Zuge der MRI wird die Frage äußerst wichtig, wer oder was als geeigneter sozialer Akteur auftreten kann. Denn ob soziale Realität aus Interaktionen entstehen kann, hängt von der Fähigkeit der betrachteten Einheit ab, alter zu sein und an symbolisch vermittelter Interaktion teilzunehmen.

In dieser Hinsicht stellen die Theorien von Goffman und Garfinkel lediglich *einen* Eckpfeiler der möglichen Untersuchungen typischer MRI Situationen dar: Wenn MRI äquivalent zu Mensch-Mensch Interaktionen (MMI) betrachtet werden, können diese Theorien zur Analyse der Situation beitragen, indem angenommen wird, dass Roboter soziale Realität mitkonstruieren. Der Vorteil bei der Anwendung dieser Theorien wäre die Adressierung der impliziten Konsequenzen in der Entwicklung sozialer Robotik, die ihrerseits explizit auf diesen Annahmen beruht. MRI analog zu MMI umzusetzen hieße nämlich nichts anderes als Roboter als gleichwertige soziale Partner in die Konstruktion dessen einzuladen, was Menschen bis dato exklusiv zustand: der sozialen Realität. Andererseits könnte die Nutzung dieser Theorien Hinweise darüber geben, in welchem Ausmaß MRI Ähnlichkeit zu MMI aufweisen sollten. Ihre Anwendung zum detaillierteren Verständnis von MRI stellt gewissermaßen eine Bewährungsprobe für den Anteil an Sozialität in der Interaktion dar.

4. Gerahmte Krisenexperimente als generische Methode

Krisenexperimente sind eine von Garfinkel entwickelte Methode, um die Grenzen dessen, was sozial ausgehandelt werden kann, offen zu legen, indem soziale Normen missachtet werden, die somit zum Vorschein treten. Befindet sich das Ausmaß der Krise innerhalb dessen, was sozial aushandelbar ist, kommen Reparaturmechanismen zur Anwendung, welche

die Störung beseitigen sollen. Ist die Interaktion zu stark irritiert, kommt es zum Abbruch, was für den beteiligten Akteur meist ein hochemotionales Ereignis darstellt (vgl. Garfinkel 1967).

Was als Krise fungieren kann und welche Mechanismen zur Reparatur geeignet sind, hängt wiederum in starkem Maße von der Rahmung der Situation ab. Rahmen sind Interpretationsschemata und helfen eine entsprechende Situation angemessen zu definieren. Sie beinhalten die Erwartungen und die erwartbaren Erwartungen des Interaktionspartners und werden bei mangelhafter Passung gewechselt.

Obgleich verschiedene Studien sich bereits der Krisenexperimente als Instrument bedient haben (teilweise auch ohne sie konkret als solche zu bezeichnen), wurde bislang noch kein systematischer Ansatz bezüglich der Etablierung dieser als allgemeiner Methodik zur Evaluation der MRI entwickelt (vgl. Weiss et al. 2008, 2010, Sirkin et al. 2015, Bauer et al. 2009, Alac et al. 2011, Muhl/Nagai 2007, Nagai/Rohlfing 2007, Short et al. 2010, Takayama et al. 2009). Krisenexperimente eignen sich aus verschiedenen Gründen in besonderem Maße zur Evaluation von MRI: Erstens, operieren sie auf einem hohen Niveau hinsichtlich des Verständnisses der Erfordernisse sozialer Handlungen. Indem sie Mensch-Mensch Interaktionen als Vergleichsmaßstab hinsichtlich der Sozialität von Handlungen setzen, ermöglichen sie auch kontrafaktische Aussagen über nicht-soziale Interaktionen. Zweitens, sind sie weniger anfällig für die häufigste Ursache von Verzerrungen: soziale Erwünschtheit. Denn diese wird in krisenhaften Situationen ebenfalls brüchig. Durch (videogestützte) Beobachtungen ist daher ein unmittelbarer Zugang zu den spontanen Reaktionen der ProbandInnen auf krisenhafte Interaktionsverläufe möglich. Diese können zudem im Nachgang durch Interviews oder Fragebögen weiter differenziert werden.

Ein dritter Vorteil der Methode liegt in ihrer kulturübergreifenden Anwendbarkeit, da das Prinzip der Methode durch Krisen Aufschlüsse über die Normalität zu erlangen unabhängig von spezifischen kulturellen Kontexten anwendbar ist. Sie operiert innerhalb der kulturell geformten Grenzen dessen, was als funktionierende Interaktion gesehen wird. Damit eignet

sie sich nicht nur in besonderem Maße für MRI Settings, sondern auch für Vergleiche zwischen Kulturen⁴, sowie Subgruppen innerhalb einer Kultur. Bezüglich der Erforschung von MRI in verschiedenen kulturellen Kontexten bedeutet dies, dass Symbole und Krisenauslöser in der Interaktion sorgfältig analysiert, transferiert und implementiert werden müssen, damit kultur- und fallsensitive Generalisierungen vorgenommen werden können (vgl. Lutze/Brandenburg 2013). Zudem sollte die Forscherin über profunde Kenntnisse der jeweiligen Kultur verfügen, um das Ausmaß von Krisen und den entsprechenden Umgang damit angemessen deuten zu können.

Typischerweise werden den Probanden bei der qualitativen Evaluation der Güte der Interaktion nach der Durchführung der Interaktionssequenz mit dem Roboter verschiedene Fragen bezüglich ihres subjektiven Empfindens dieser Erfahrung gestellt. Verglichen mit den altbekannten Mensch-Mensch Interaktionen (MMI) werden MRI häufig als enttäuschend (obgleich ähnlich) wahrgenommen. Die Interaktionssequenz wird überwiegend vom Menschen ausgeführt und bestimmt. Der Mensch füllt die Lücke, die aufgrund der Unfähigkeit des Roboters entsteht. In MRI Experimenten führt diese spezielle – wenngleich typische – Situation häufig zu einer positiveren Bewertung der erfahrenen Güte der Interaktion. Die Probanden werden vermutlich ihre unternommenen Anstrengungen betonen, den Interaktionsfluss aufrecht zu erhalten und das hervorheben, was sie als Erwartungen des Forschers wahrgenommen haben. Als Methode liefern Krisenexperimente eine authentischere Reaktion, da neben den subjektiven Einschätzungen der Probanden selbst, die beobachtbaren Reparaturmechanismen bzw. den Abbruch als Indiz für die Qualität der Interaktion herangezogen werden können. Diese lassen sich dann bspw. nach Art, Zeitpunkt und Intensität klassifizieren und mit anderen qualitativen Daten triangulieren.

4 Vergleichende Studien, die die Entwicklung von Sozialer Robotik in Europa und Japan analysieren, kommen zu dem Schluss, dass diese sich soziologisch betrachtet hinsichtlich des Verständnisses von Roboter-Agency und des Konzeptes einer angemessenen Nutzer-Roboter Interaktion stark unterscheiden (vgl. Wagner 2013, MacDorman et al. 2009, Kaplan 2004). In Europa steht die Autonomie des Roboters im Vordergrund, während in Japan durch vordefinierte und standardisierte MRI-Sequenzen diese gerade außer Kraft gesetzt werden soll (vgl. Matsuzaki/Lindemann 2015). Gerahmte Krisenexperimente können in beiden Settings durchgeführt werden.

Zudem ist der Blick auf den Kontext für MRI Situationen – mit oder ohne Krisenexperimente – unerlässlich. Wird die MRI durch den Kontext als Experiment gerahmt, kommt dem Roboter die Eigenschaft einer Maschine zu, die es zu testen gilt. Das Scheitern der Interaktion wäre in diesem Fall keine Angelegenheit von höherer Bedeutung. Durch Krisenexperimente in Laborkontexten und den umsichtigen Umgang mit Rahmungen kann, dieser Mechanismus jedoch durchbrochen werden und wertvolle Erkenntnisse hinsichtlich (sozialer) Mensch-Roboter Interaktion generieren. Wird der Roboter hingegen in alltäglichen Situationen eingeführt, wo der Kontext mehrdeutig ist und die Menschen (womöglich beeinflusst vom Bild von Robotern aus den Medien – Filmen, Science Fiction Literatur, etc.) nicht genau wissen, über welche Fähigkeiten der Roboter verfügt, entsteht eine höchst ungewisse Situation. Mensch-Roboter Interaktionen scheitern hier durch unklare Erwartungen deutlich schneller. Der Blick auf die Rahmung weist auch auf einen wichtigen Aspekt hin, der mit dem in der Robotikforschung viel diskutierten „Uncanny Valley“⁵ zu tun hat. Ist der Status des Roboters uneindeutig; weiß ego also nicht, ob der Roboter nur eine Maschine ist oder ob er tatsächlich menschenähnliche Fähigkeiten besitzt, so wird die Interaktion zutiefst gestört. Die Erwartungen an den Roboter variieren mit der Situationsdefinition.

5. Vergleich zweier empirischer Fallstudien

Die bislang vorgestellten Überlegungen zu gerahmten Krisenexperimenten sollen nun anhand zweier empirischer Fallbeispiele diskutiert werden. Eine der Studien stützt sich auf Befunde aus einem Laborkrisenexperiment und identifiziert verschiedene Strategien, wie die Probanden mit der induzierten Krise umgehen (vgl. Muhl/Nagai 2007). Die zweite berichtet hingegen über Forschung in einer Alltagssituation mit kognitiv orientierten und desorientierten älteren Bewohnern einer Pflegeeinrichtung und gibt wichtige Hinweise für unterschiedliche Rahmungen (vgl. Compagna/Muhl 2012). Diese empirischen Arbeiten innerhalb eines ähnlichen theoretischen Rahmens zeigen das Potential unserer Methode, fall-sensitive Schlüsselfaktoren innerhalb des breiten Spektrums an MRI-Situationen zu identifizieren.

5 Das Phänomen des „Uncanny Valley“ (unheimliches Tal) beschreibt den Umstand, dass die Akzeptanz von Robotern zunächst mit zunehmender anthropomorpher Gestalt ansteigt, dann jedoch schlagartig abfällt und erst ab einem sehr hohen Grad an Ähnlichkeit wieder ansteigt.

Fall 1

Muhl & Nagai (2007) zeigen in ihrer Studie, dass mit Krisenexperimenten auch in Laborkontexten beachtliche Resultate generiert werden können. Ohne ihr Experimentaldesign theoretisch zu reflektieren, nutzten sie eine typische Täuschungsstrategie. Die Probanden wurden instruiert, einem Roboter einige einfache Aufgaben beizubringen. Der Roboter war in diesem Fall ein Babygesicht auf einem Bildschirm (vgl. Ogino et al. 2006). Augen, Augenlider, Augenbrauen und der Mund waren animiert. Der Roboter war mit einem biologisch inspirierten Salienz-Mechanismus ausgestattet (vgl. Nagai et al. 2006). Sein Blick folgte den wichtigsten Merkmalen der Szenerie und signalisierte seinem menschlichen Interaktionspartner somit Aufmerksamkeit bzw. forderte diese ein. Ein Sprachverarbeitungssystem oder akustische Sensoren waren nicht vorgesehen (vgl. Muhl/Nagai 2007). Die induzierte Krise bestand darin, den Blick des Roboters abdriften zu lassen, woraufhin die Erwartungen der Probanden irritiert wurden und diese ihrerseits verschiedene Reparaturmechanismen zur Wiedererlangung der Aufmerksamkeit anwandten. Diese umfassten bspw.:

- den Blick in Richtung dessen, worauf der Roboter schaut
- den Versuch, die Aufmerksamkeit des Roboters wieder auf sich zu ziehen (z.B. durch lauterer Sprechen oder größere Bewegungen)
- den Versuch, die Aufmerksamkeit des Roboters auf das Objekt zu lenken
- den Versuch, in den Aufmerksamkeitsbereich des Roboters zu kommen (vgl. Muhl/Nagai 2007).

Obgleich die Interaktion rudimentär war, zeigen die angewandten Reparaturstrategien, dass die Probanden in gewissem Maße Interaktionsfähigkeiten des Roboters erwarteten. Die kognitive Rahmung ego's hinsichtlich seines Interaktionspartners alter ist für das Gesamturteil über alter und damit auch für die Wahl der geeigneten Vorgehensweise entscheidend. In diesem Zusammenhang könnte es mit einer sorgfältig durchgeführten Rahmenanalyse gelingen, wichtige Faktoren, die für den Gesamtergebnis des MRI-Experimentes primär verantwortlich sind, zu identifizieren. Die jeweilige Situationsdefinition des Probanden bestimmt, welche bzw. ob überhaupt Reparaturmechanismen angewandt werden. Durch die

Berücksichtigung der Rahmung kann dann abgeschätzt werden, wie diese zu bewerten sind.

Fall 2

Die zweite empirische Studie wurde in einer Pflegeeinrichtung für ältere Menschen in einem Alltagssetting durchgeführt. Ein Serviceroboter hatte die Aufgabe ein Glas Wasser an die Einwohner der Pflegeeinrichtung zu servieren (vgl. Compagna 2010, 2012). Er sollte die Bestellung der Menschen annehmen und das Glas dann an die richtige Person übergeben und diese verbal adressieren. Die BewohnerInnen antworteten jedoch oftmals nicht, da sie es vorzogen, die anderen anwesenden Personen zu adressieren. In den Fällen, in denen die Einwohner das Getränk annahmen, sollte der Roboter ihnen dafür danken, was seitens der Einwohner ebenfalls meistens ignoriert wurde. Der Roboter war nicht in der Lage flexibel mit der Situation umzugehen und auf die Zurückweisung mit einer Aufforderung zu reagieren, bspw. in dem er sie in irgendeiner Weise kommentiert, was die Möglichkeit einer erfolgreichen Kommunikation erhöht hätte. Wenn eine erwartete Handlung des Roboters nicht eintritt, wurde die Interaktion abgebrochen bzw. auf andere anwesende Menschen fokussiert. Die Bewohner versuchten gar nicht erst, ein gemeinsames Verständnis mit dem Roboter herzustellen. In diesen Fällen ist Interaktion im soziologischen Sinne nicht nur durch ihr Scheitern gefährdet – sie findet gar nicht statt. Ein besonderer Fall ließ sich in ebendieser Studie bei einer Interaktionssequenz mit einer demenziell erkrankten Probandin beobachten. Die Interaktion zwischen ihr und dem Roboter verlief erfolgreich und völlig reibungslos – die Versuchsperson hat den Roboter als vollwertigen sozialen Akteur und Interaktionspartner betrachtet, obgleich dieser nicht in der Lage war zu reagieren bzw. die Reaktion für die anderen Beteiligten nicht der zu erwartenden entsprach. Ähnlich positive Reaktionen auf Roboter zeigen sich auch in anderen Studien mit demenziell Erkrankten (vgl. Giusti/Marti 2006, Broekens et al. 2009).

Vergleich

Der Vergleich der beiden Studien bringt wichtige Erkenntnisse hinsichtlich der Rahmung: Im experimentellen Setting sind die Versuchspersonen äußerst motiviert, die Interaktion zu reparieren und unternehmen die entsprechenden Anstrengungen. Erst nach vielen erfolglosen

Versuchen brechen sie die Interaktion ab. Die Alltagssituation im Pflegeheim zeigt hingegen ein ganz anderes Bild: Die Versuchspersonen unternehmen keinerlei Anstrengungen, in Interaktion mit dem Roboter zu treten. Sie ist damit schlichtweg gescheitert. Alltagssituationen stellen für Interaktionsexperimente mit Robotersystemen den strengsten Test dar – wenn hier induzierte Krisen repariert werden, kann die Interaktion als sozial bezeichnet werden. Laborkontexte sind hier weniger streng, da die Versuchspersonen sich dazu verpflichtet fühlen, die ihnen gestellte Aufgabe angemessen zu erfüllen und ein Abbruch so unwahrscheinlicher wird. Der Vergleich zwischen kognitiv orientierten und desorientierten Patienten der Pflegeeinrichtung weist wiederum auf die subjektiv unterschiedliche Rahmung hin. In beiden Fällen waren keine Reparaturstrategien beobachtbar, da für die einen keine Interaktion stattfand, die reparaturbedürftig gewesen wäre (Abbruch), während die anderen sich voll und ganz auf eine Interaktion einließen, die subjektiv erfolgreich verlief und damit ebenso wenig reparaturbedürftig war. Im Falle der kognitiv orientierten Patienten zeigte sich zudem die entscheidende Rolle anderer anwesender Personen, die im Falle einer Störung adressiert wurden, statt die Interaktion zu reparieren.

6. Zusammenfassung und kritische Reflexion

Eine Methode, die auf Krisenexperimenten als zentralem Instrument basiert und gleichzeitig die Bedeutung der Rahmung betont, ist den zitierten Befunden nach zu urteilen höchst geeignet, um reliable Ergebnisse hinsichtlich der Qualität der Interaktion zwischen einem Menschen und einem Roboter aus genuin soziologischer Perspektive zu erbringen. Die Beobachtung *ob* und *wie* eine (explizit durch den Forscher) induzierte Krise durch den menschlichen Akteur repariert wird kann zur bedeutungsvollen Evaluation einer MRI herangezogen werden und eignet sich zudem zur Identifikation von Unterschieden zwischen einzelnen Individuen (vgl. Muhl/Nagai 2007, Compagna/Muhl 2012).

Die Reflektion der Rahmung als wichtiger Einflussvariable ist in zweifacher Hinsicht entscheidend: einerseits für die Entscheidung, welcher Rahmen für einen Versuchsaufbau gewählt werden sollte und andererseits für die sorgfältige Interpretation der Ergebnisse. Wird der Rahmen überlegt gewählt, können die Ergebnisse für die Evaluation der Qualität der MRI äußerst hilfreich sein. Wenn der menschliche Akteur Strategien zur Reparatur der Interaktion mit dem Roboter anwendet, weist diese insofern basale soziale Züge auf, da er

dann – zumindest einseitig – seine Erwartungen revidiert und sie an die vermuteten Erwartungen des Roboters anpasst. Selbst wenn der menschliche Akteur sich darüber bewusst ist, dass der Roboter eine Maschine ist, die nicht in der Lage ist, den Ablauf der Interaktion ihrerseits zu reparieren (man könnte auch sagen, dass der Roboter nicht in der Lage ist, doppelte Kontingenz zu prozessieren bzw. auf dieser Basis einer Hypothese folgt, den Interaktionsfluss aufrechtzuerhalten oder wiederherzustellen), kann der menschliche Akteur den Roboter dennoch als Entität betrachten, die als sozialer Akteur behandelt werden kann, solange dieser basale interaktive Fähigkeiten besitzt. Es entsteht eine Form von Sozialität, die sich zwar von der zwischen Menschen unterscheidet, aber gleichzeitig mehr ist, als bloße Mensch-Maschine Interaktion. Diesem Andersartigen und den Eigenschaften des Roboters, die beim Menschen derart soziale Reaktionen hervorrufen, gilt es in künftiger MRI Forschung weiter nachzugehen.

Zusammenfassend lässt sich sagen, dass wenn die Rahmung berücksichtigt und richtig gewählt wird, die Art und Weise, wie Reparaturstrategien seitens des menschlichen Akteurs gewählt werden (bspw. kombiniert mit der beobachteten Häufigkeit, Anzahl und Dauer), als Hinweis für die Qualität der MRI genutzt werden können. Auf dieser Grundlage können Forscher nützliche Informationen für die Weiterentwicklung sozialer Roboter und ihrer Interaktionsfähigkeiten erlangen.

Literatur

Alac, Morana / Javier Movellan / Fumihide Tanaka (2011): When a Robot Is Social: Spatial Arrangements and Multimodal Semiotic Engagement in the Practice of Social Robotics. *Social Studies of Science*, 0306312711420565.

Baecker, Dirk (2011): Who qualifies for communication? A systems perspective on human and other possibly intelligent beings taking part in the next society. In: *Technikfolgenabschätzung - Theorie und Praxis* 20 (1), pp. 17-26.

Bauer, Andrea / Klaas Klasing / Georgios Lidoris / Quirin Mühlbauer / Florian Rohrmüller / Stefan Sosnowski / Tingting Xu / Kolja Kühnlenz / Dirk Wollherr / Martin Buss (2009): The Autonomous City Explorer: Towards Natural Human-Robot Interaction in Urban Environments. In: *International Journal of Social Robotics* 1, no. 2, pp. 127–40.

Berger, Peter L. / Luckmann, Thomas (1966/2003): *Die gesellschaftliche Konstruktion der Wirklichkeit - Eine Theorie der Wissenssoziologie* (19. Aufl.). Frankfurt am Main: Fischer-Taschenbuch-Verlag.

Broekens, J., Heerink, M., & Rosendal, H. (2009): Assistive social robots in elderly care: a review. In: *Gerontechnology*, 8, no. 2, pp. 94–103.

- Burghart, Catherina / Roger Haeussling (2005): Evaluation Criteria for Human Robot Interaction. In: Proceedings of the Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction, pp. 23–31.
- Compagna, Diego (2010): Reconfiguring the User: Raising Concerns over User-Centered Innovation. In: Proceedings ECAP10. VIII European Conference on Computing and Philosophy, pp. 332–336.
- Compagna, Diego (2012): Lost in Translation? The Dilemma of Alignment within Participatory Technology Developments. In: Poiesis & Praxis, 9 (1-2), pp. 125–143.
- Compagna, Diego / Muhl, Claudia (2012): Mensch-Roboter-Interaktion – Status der technischen Entität, Kognitive (Des)Orientierung und Emergenzfunktion des Dritten, in: „Muster und Verläufe der Mensch-Technik-Interaktivität“, Band zum gleichnamigen Workshop am 17./18. Juni 2011 in Berlin, Technical University Technology Studies, Working Papers, TUTS-WP-2-2012, Julian Stubbe, Mandy Töppel (eds.), Berlin, 19–34.
- Feil-Seifer, David / Kristine Skinner / Maja J Matari (2007): Benchmarks for Evaluating Socially Assistive Robotics. In: Interaction Studies 8, no. 3, pp.423–439.
- Garfinkel, Harold (1967): Studies in Ethnomethodology. Cambridge: Polity Press.
- Giusti, L., & Marti, P. (2006): Interpretative dynamics in human robot interaction. In: Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on. IEEE, pp. 111–116.
- Goffman, Erving (1963): Behavior in Public Places. New York: Free Press.
- Goffman, Erving (1967): Interaction Ritual: Essays on Face-to-Face Behavior. New York: Anchor Books.
- Goffman, Erving (1971): Relations in Public: Microstudies of the Public Order. New York: Basic Books.
- Goffman, Erving (1974): Frame Analysis. New York: Harper & Row.
- Hahn, Alois (2004): Der Mensch in der deutschen Systemtheorie. In: Bröckling, Ulrich / Paul, Axel T. / Kaufmann, Stefan (Hg.): Vernunft - Entwicklung - Leben. Schlüsselbegriffe der Moderne. Festschrift für Wolfgang Eßbach. München: Fink, pp. 279–290.
- Herrmann, Guido (eds.) (2013): Social Robotics - 5th International Conference, ICSR 2013, Bristol, UK, October 27 - 29, 2013 ; Proceedings. Lecture Notes in Computer Science 8239. Cham: Springer.
- Kaplan, Frédéric (2004): Who Is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots. In: International Journal of Humanoid Robotics 1, no. 03, pp. 465–480.
- Kahn Jr, P. H., Ishiguro, H., Friedman, B., Kanda, T., Freier, N. G., Severson, R. L., & Miller, J. (2007): What is a Human? Toward psychological benchmarks in the field of human–robot interaction. In: Interaction Studies, 8 (3), pp. 363–390.
- Lindemann, Gesa (1999): Doppelte Kontingenz und reflexive Anthropologie. In: Zeitschrift für Soziologie 28 (3), pp. 165–181.
- Luhmann, Niklas (1984): Soziale Systeme. Grundriß einer allgemeinen Theorie. Frankfurt a.M.: Suhrkamp.
- Lutze, Maxie / Brandenburg, M. (2013). Do we need a new Internet for elderly people? A crosscultural investigation. In P.L.P. Rau (Ed.): CCD/HCI 2013, Part II, LNCS 8024, pp. 441–450, Springer-Verlag Berlin Heidelberg

- MacDorman, Karl F / Sandosh K Vasudevan / Chin-Chang Ho (2009): Does Japan Really Have Robot Mania? Comparing Attitudes by Implicit and Explicit Measures. In: *AI & Society* 23, no. 4, pp. 485–510.
- Matsuzaki, Hironori / Lindemann, Gesa (2015 – Online First): The autonomy-safety-paradox of service robotics in Europe and Japan: a comparative analysis. In: *AI & Society* 31, no. 4, pp. 501-517.
- Mead, G. H. (1934/2013): *Geist, Identität und Gesellschaft: aus der Sicht des Sozialbehaviorismus* (17. Aufl.). Frankfurt am Main: Suhrkamp.
- Muhl, Claudia / Nagai, Yukie (2007): Does disturbance discourage people from communicating with a robot? 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Jeju, Korea.
- Nagai, Yukie / Asada, Monoru / Hosoda, Koh (2006): Learning for joint attention helped by functional development. In: *Advanced Robotics*, 20 (10), pp. 1165-1181.
- Nagai, Yukie / Katharina J. Rohlfing (2007): Can Motionese Tell Infants and Robots ‘What to Imitate’? In: *Proceedings of the 4th International Symposium on Imitation in Animals and Artifacts*, pp. 299–306.
- Ogino, Masaki / Watanabe, Ayako / Asada, Monoru (2006): Mapping from facial expression to internal state based on intuitive parenting. In: *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, pp. 182-183
- Pitsch, K., & Wrede, S. (2014). When a robot orients visitors to an exhibit. Referential practices and interactional dynamics in real world HRI. In *Ro-Man 2014* (pp. 36-42)
- Pitsch, K., Lohan, K. S., Rohlfing, K., Saunders, J., Nehaniv, C. L., & Wrede, B. (2012). Better be reactive at the beginning. Implications of the first seconds of an encounter for the tutoring style in human-robot-interaction. In *Ro-Man 2012* (pp. 974-981). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6343876&tag
- Pitsch, K., Vollmer, A. -L., & Mühlig, M. (2013). Robot feedback shapes the tutor’s presentation. How a robot’s online gaze strategies lead to micro-adaptation of the human’s conduct. *Interaction Studies*, 14(2), 268-296. doi:<http://dx.doi.org/10.1075/is.14.2.06pi>
- Short, Elaine / Justin Hart / Michelle Vu / Brian Scassellati (2010): No Fair an Interaction with a Cheating Robot. In: *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, pp. 219–26.
- Sirkin, David / Brian Mok / Stephen Yang / Wendy Ju (2015): Mechanical Ottoman: How Robotic Furniture Offers and Withdraws Support. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 11–18. ACM.
- Takayama, Leila / Victoria Groom / Clifford Nass: I’m Sorry, Dave (2009): I’m Afraid I Won’t Do That: Social Aspects of Human-Agent Conflict. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2099–2108.
- VDI (2012): VDI Guideline 6220, Part 1: Biomimetics – Conception and Strategy. Differences between Biomimetic and Conventional Methods/Products. Verein Deutscher Ingenieure e.V., Düsseldorf.
- VDI (2013): VDI Guideline 6222, Part 1: Biomimetics – Biomimetic Robots. Verein Deutscher Ingenieure e.V., Düsseldorf.
- Wagner, Cosima (2013): *Robotopia Nipponica: Recherchen zur Akzeptanz von Robotern in Japan*. Marburg: Tectum-Verl.

Weiss, Astrid / Judith Igelsböck / Manfred Tscheligi / Andrea Bauer / Kolja Kühnlenz / Dirk Wollherr / Martin Buss (2010): Robots Asking for Directions: The Willingness of Passers-by to Support Robots. In: Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, 23–30. IEEE Press.

Weiss, Astrid / Regina Bernhaupt / Manfred Tscheligi / Dirk Wollherr / K Kühnlenz / Martin Buss (2008): A Methodological Variation for Acceptance Evaluation of Human-Robot Interaction. In: Public Places.” In Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on, 713–18. IEEE.

Wrede, B., Rohlfing, K. J., Hanheide, M., & Sagerer, G. (2009): Towards learning by interacting. In: B. Sendhoff, E. Körner, O. Sporns, H. Ritter, & K. Doya (Hrsg.): Creating Brain-Like Intelligence - From Basic Principles to Complex Intelligent Systems. Springer, pp. 139–150.

Wykowska, Agnieszka / Ryad Chellali / Md Mamun Al-Amin / Hermann J Müller (2014): Implications of Robot Actions for Human Perception. How Do We Represent Actions of the Observed Robots? In: International Journal of Social Robotics 6, no. 3, pp. 357–66.